

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 63-168774
 (43)Date of publication of application : 12.07.1988

(51)Int.Cl.

G06F 15/38

(21)Application number : 62-000709

(71)Applicant : FUJITSU LTD

(22)Date of filing : 06.01.1987

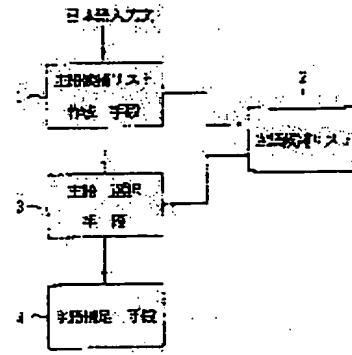
(72)Inventor : NAKADA TERUO
UCHIDA YOSHIKI

(54) SUBJECT SUPPLEMENTING SYSTEM FOR MECHANICAL TRANSLATION SYSTEM

(57)Abstract:

PURPOSE: To obtain a natural translation by providing a subject candidate list in order to supplement the subject of a Japanese sentence from which the subject is omitted, selecting the most suitable one therefrom and supplementing as the subject and translating.

CONSTITUTION: At the time of reaching the sentence from which the subject is omitted, a subject selecting means 3 sequentially watches the subject candidate list 2 from the last part and selects one which can be the subject grammatically and in a sense. Therefore, the attribute of the reference of the selection as well as a word is stored in the subject candidate list 2. Ordinarily, since the subject of the sentence in which the subject is omitted is a lately appearing noun, when the most lately appearing noun of the nouns grammatically correctly connected nouns is used as the subject of the sentence in which the subject is not present, the correct translation can be formed. When the proper subject candidate is not found, or when there are more than two candidates and which is more suitable cannot be decided, the translation is executed as by using a passive voice.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

⑯ 公開特許公報 (A)

昭63-168774

⑯ Int.Cl.
G 06 F 15/38識別記号
7313-5B

⑯ 公開 昭和63年(1988)7月12日

審査請求 未請求 発明の数 1 (全5頁)

⑯ 発明の名称 機械翻訳システムにおける主語の補足方式

⑯ 特願 昭62-709

⑯ 出願 昭62(1987)1月6日

⑯ 発明者 中田輝生 神奈川県川崎市中原区上小田中1015番地
内 富士通株式会社

⑯ 発明者 内田好昭 神奈川県川崎市中原区上小田中1015番地
内 富士通株式会社

⑯ 出願人 富士通株式会社 神奈川県川崎市中原区上小田中1015番地

⑯ 代理人 弁理士 井桁貞一

明細書

1. 発明の名称

機械翻訳システムにおける主語の補足方式

2. 特許請求の範囲

(1) 日本語を原言語とする機械翻訳システムにおいて、

入力文を解析して名詞を抽出し、主語候補として格納し、リストとして管理する主語候補リスト作成手段(1)と、

主語候補リスト作成手段(1)の抽出した新しい主語候補を次々とつないで格納した主語候補リスト(2)と、

主語の省略された文に到達したとき主語候補リスト(2)の後の方から主語となり得るものを選択する主語選択手段(3)と、

主語選択手段(3)の選択した名詞の代名詞を、前記主語なし文の主語として補足する主語補足手段(4)を備え、

主語を補足された文により翻訳を実行するよ

う構成したことを特徴とする機械翻訳システムにおける主語の補足方式。

(2) 上記主語選択手段(3)による主語の選択が不可能なとき、目的語を主語とする受動体として翻訳を行うよう構成したことを特徴とする特許請求の範囲第1項記載の機械翻訳システムにおける主語の補足方式。

(3) 上記主語選択手段(3)による主語の選択法を、文書の種類に応じて切り換える手段を備えるよう構成したことを特徴とする特許請求の範囲第1項記載の機械翻訳システムにおける主語の補足方式。

3. 発明の詳細な説明

[概要]

機械翻訳システム、特に日本語のように主語が存在しない場合が多い原言語から行う機械翻訳システムにおいて、主語を省略された日本語文の主語を補足するため、主語候補リストを備え、この中から最も確からしいものを選択し、主語として

補足して翻訳するようにしたものであり、これによってより自然な訳文が得られる。

[産業上の利用分野]

本発明は、日本語からの機械翻訳に係わり、原文で省略された主語を補う方式に関する。

日本語においては、主語が省略されることが多い。一方、欧米の言語では主語は必須であるから、何等かの操作が必要とされる。

[従来の技術]

従来、日本語からの機械翻訳において、主語なし文に対しては、目的語を主語とする文（受動体）を生成することにより、翻訳を成功させていた。

第4図は、従来技術による機械翻訳の例を示す図である。

日本語例文に示すように、最初の文aには主語があるが、これに続く文b, c, dには主語が省略されている。

これらに対する従来技術による翻訳は、訳例1

に示すようになる。即ち、主語なし文b, c, dは、目的語を主語として受動体として生成されている。

このように、従来方法による機械翻訳では、翻訳としては誤っていないけれども、不自然な、ぎごちない文が生成されてしまう。

[発明が解決しようとする問題点]

上記に説明のように、従来の方式によれば、主語無し文の翻訳が不自然なぎごちない文になってしまふものであった。

そこで、省略された主語を補うことが要求されるが、誤った主語を補ってしまうと、翻訳の精度を落すことになる。

本発明は、このような従来の問題点を解消した新規な主語の補足方式を提供しようとするものである。

[問題点を解決するための手段]

第1図は本発明の日本語を原言語とする機械翻

訳システムにおける主語の補足方式の原理ブロック図を示す。

図において、1は主語候補リスト作成手段であり、入力文を解析し、名詞を抽出し、主語候補として主語候補リスト2に格納する。主語候補リスト2は、次々と新しく現れた主語候補をつないで格納する。

3は主語選択手段であり、主語の省略された文に到達したとき、主語候補リスト2の要素をリストの後の方からそれが主語となり得るかを調べ、選択する。

4は主語補足手段であり、主語選択手段3の選択した名詞の代名詞を主語なし文の主語として補足する。

[作用]

主語候補リスト作成手段は、入力された日本語文を解析して名詞を抽出し、以後の文の主語になる可能性のあるものとして、主語候補リストの後に次々とつないでいく。

主語の省略された文に到達したとき、主語選択手段は、主語候補リストを最後の方から順に見ていて、文法的に、意味的に主語になり得るものを選択する。そのため、主語候補リストには、語のはかに選択の参考となる属性を記載する。

通常、主語省略の文の主語は、最近出てきた名詞であるから、文法的に正しくつながる名詞のうち、最も最近現れた名詞を主語無し文の主語として採用すれば、正しい訳文ができると考えられる。

適当な主語候補が見つからないとき、または二つ以上の候補があり、どれが適当か判断ができないときは、受動体として翻訳を行う。

文書の種類によっては、特に主語になり易い語がある。例えば、物語文では、その主人公など、また論文では、それに関する専門用語（物質名、実験装置等）を優先するのが適当である。

従って、文書の種類に応じて、選択方法を切り換える手段を備えて、これに対応することができる。

【実施例】

以下第2図および第3図に示す実施例により、本発明をさらに具体的に説明する。

第2図は、本発明の一実施例における主語候補リストの記述例を示す図である。

図に示すように、この主語候補リストには、①主語候補の語に加えて、②その語の属性、③その代名詞、および④参照回数カウンタを付随させて格納する。

属性は、文法的に正しい主語になり得るかを調べる際に使用する。

代名詞は、この語が主語なし文の主語として選択されたとき、この代名詞を主語として文を生成する。

参照回数カウンタは、主語なし文の主語として採用された回数を計数し、主語候補リストの容量制限により古いものから捨てる際に、残すための条件とする。

図において、上側の矢印一は、新しい語がつながれていくことを示し、下側の矢印一は、このリ

ストで次に参照する語のポインタを示し、新しいものから順次参照し、nil はこれ以上参照する語はないことを示す。

以下、第4図に示した日本語例文について、本実施例による動作を説明する。

(1)まず、最初の文 a の翻訳時に、主語の「チョウ チョウ」と目的語の「(みかんの)木」が主語候補リストに追加される。「チョウ チョウ」(butterfly) には属性として動物(animal)が、代名詞として it が記載され、「木」(tree)には属性として植物(plant)が、代名詞として it が記載され、「チョウ チョウ」は主語であったから参照回数カウンタは 1 とする。

次に、文 b の翻訳においては、「(何か) つけて・・」の主語が省略されているのが検出される。主語選択手段は、「つける」は有意思動詞であるから、主語は有意思体であると推定し、主語候補リストを後から、動物か人間乃至擬人化され易いものという条件で検索する。「木」は不適であるが、「チョウ チョウ」は残り、主語候補リストに登録された語が他に無いので「チョウ チョウ」が文 b の主語として選択される。主語候補リスト内の「チョウ チョウ」の参照回数カウンタを 1 だけ増す。

(3)次に、文 c の翻訳において、「生んでいる」の主語省略されているのが検出される。主語選択手段は、「生んでいる」は有意思動詞であるから、主語は有意思体であると推定し、主語候補リストを後から、動物か人間乃至擬人化され易いものという条件で検索する。「葉」、「おなか」、「木」は不適であるが、「チョウ チョウ」は残り、主語候補リストに登録された語が他に無いので「チョウ チョウ」が文 c の主語として選択される。主語候補リスト内の「チョウ チョウ」の参照回数カウンタを 1 だけ増す。

(5)次に、文 d の翻訳において、「卵」(egg) が抽出され、その属性、代名詞と共に、主語候補リストに追加される。このとき、候補リストの長さが "4" に設定されているならば、候補リストの古い方(より前に現れた語)で、一度も参照されていないもの、「木」(tree)を捨てる。一度も参照されていないものが無いならば、最も古いものを捨てる。

(7)主語選択手段において、主語候補リストから適

語候補リストに登録された語が他に無いので、「チョウ チョウ」が文 b の主語として選択される。主語候補リスト内の語「チョウ チョウ」が主語として使用されたことを示す参照回数カウンタを 1 だけ増す。

(3)文 b の名詞が抽出され、「おなか」(stomach) および「葉」(leaf) がそれぞれの属性: 肉体(body)、植物(plant) ならびに代名詞 it と共に、主語候補リストに追加される。

(4)次に、文 c の翻訳では、「(何を) している」の主語が省略されているのが検出される。主語選択手段は、「している」は有意思動詞であるから、主語は有意思体であると推定し、主語候補リストを後から、動物か人間乃至擬人化され易いものという条件で検索する。「葉」、「おなか」、「木」は不適であるが、「チョウ チョウ」は残り、主語候補リストに登録された語が他に無いので「チョウ チョウ」が文 c の主語として選択される。主語候補リスト内の語「チョウ チョウ」の参照回数カウンタを 1 だけ増す。

当な候補が見つからないとき、または二つ以上の候補があり、どれが適当か判断できないときは、選択不能を出力し、受動体として翻訳を行う。

第3図は、このようにして主語を補足し、英語に翻訳した文を訳例2として示す。

第4図に示した従来技術による訳例1に比較して、極めて自然な訳文となっていることが判る。

【発明の効果】

以上説明のように本発明によれば、日本語からの機械翻訳システムにおいて、省略された主語として最も確からしいものを選択し、補足することにより翻訳精度を向上することができ、その実用上の効果は極めて大である。

4. 図面の簡単な説明

第1図は本発明の原理ブロック図、

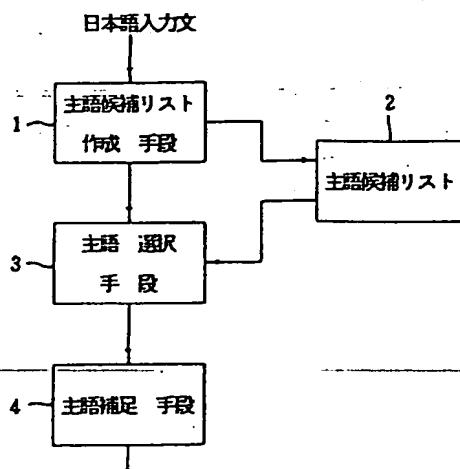
第2図は本発明の一実施例による主語候補リストを示す図、

第3図は本発明の一実施例による翻訳例を示す図、

第4図は従来技術による翻訳例を示す図である。図面において、

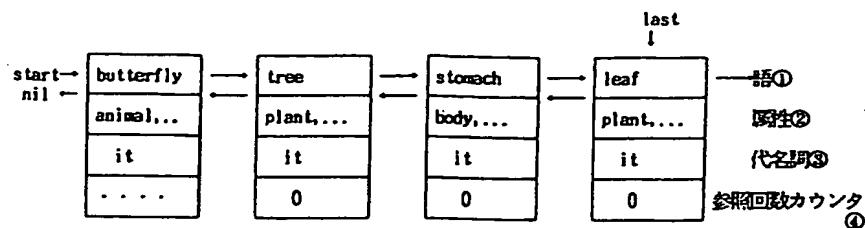
- 1は主語候補リスト作成手段、
 - 2は主語候補リスト、
 - 3は主語選択手段、
 - 4は主語補足手段、
- をそれぞれ示す。

代理人 弁理士 井桁 貞一



本発明の原理ブロック図

第1図



本発明の一実施例における主語候補リストを示す図

第 2 図

訳 例 2

- a. A butterfly came flying to an orange tree.
- b. Sometimes, bending the stomach, it put something on the leaf.
- c. What is it doing ?
- d. It is bearing the egg.

本発明の一実施例による翻訳例を示す図

第 3 図

日本語例文

- a. チョウウツウがみかんの木に飛んできました。
- b. 時々、おなかを曲げて、葉に何かつけています。
- c. 何をしているのでしょうか。
- d. 卵を生んでいます。

訳 例 1

- a. A butterfly came to an orange tree.
- b. Sometimes, the stomach being bent, something is put on the leaf.
- c. What is being done ?
- d. The egg is being born.

従来技術による翻訳例を示す図

第 4 図